Structural Studies Involving Different HIV-1 V3 Loops

Paolo Catasti and Goutam Gupta

MS K710, Los Alamos National Laboratory, Los Alamos, New Mexico 87545

Introduction

Studies on the feasibility of a subunit vaccine to protect against human immunodeficiency virus (HIV) infection have principally focused on the third variable (V3) loop of the envelope surface protein. One of the neutralizing determinants of HIV-1 is located inside the V3 loop. However, progress toward a vaccine based on neutralizing determinants has been impeded by the amino acid sequence variability in the V3 loop of different HIV isolates. The elusive nature of the V3 loop structure prompted us to carry out a systematic study on different isolates in an attempt to identify a common structural motif in the V3 loop regardless of the amino acid sequence variability. We have performed 2D NMR structural studies on three different V3 loop peptides: MN, Haiti (Haiti 6004; L07201), and RF (Catasti *et al*., 1995 & 1996). The three V3 loops were all 35 residues long and S-S bridged at the terminals. The NMR studies were carried out first in water, then in a 70%/30% mixture of water/trifluoroethanol l (TFE). TFE is a solvent widely used in NMR on peptides, for its property to unmask helical propensities of hydorphobic residues.

Figure 1 shows that similar secondary structures are observed for the three different V3 loops: a GPG(K/R) crest in the center of the neutralizing determinant, two extended regions flanking the central crest, and a helical region in the C-terminal domain observed only in the water/TFE mixture. The RF V3 peptide did not dissolve in the water/TFE mixture, therefore we could run the experiments only in an aqueous solution. Structural prediction studies revealed that the variability in sequence and structure of the V3 loop is confined to the N and C-terminal side of the conserved GPG crest. Figure 2 is a summary of the NMR secondary structural assignments (Catasti *et al*., 1995 & 1996), and the results of several secondary prediction algorithms. With the exception of the PSA method, most of the algorithms fail to identify the alpha helix in the C-terminal portion of the V3 loops.

References

Catasti, P., Bradbury, E.M., and Gupta G. (1996) Structure and polymorphism of HIV-1 third variable loops. *J. Biol. Chem*., **271**(14):8236–8242.

Catasti, P., Fontenot, J.D., Bradbury, E.M., and Gupta G. (1995) Local and global structural properties of the HIV-MN V3 loop. *J. Biol. Chem*., **270**(5):2224–2232.

Deleage, G., and Roux, B. (1987) An algorithm for protein secondary strcuture prediction based on class prediction. *Prot. Eng*., **1**:289–294.

Geourjon, C., and Deleage, G. (1994) SOPM : a self optimised prediction method for protein secondary structure prediction. *Prot. Eng*., **7**:157–164.

Geourjon, C., and Deleage, G. (1995) SOPMA : Significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *CABIOS*, **11**:681–684

Gibrat, J.F., Garnier, J., and Robson, B. (1987) Further developments of protein secondary structure prediction using information theory. New parameters and consideration of residue pairs. *J. Mol. Biol*., **198**(3):425–443.

King, R.D., and Sternberg, M.J.E. (1996) Identification and application of the concepts important for accurate and reliable protein secondary structure prediction. *Prot. Science*, in press.

Kneller, D.G., Cohen, F.E., and Langridge, R. (1990) Improvements in Protein Secondary Structure Prediction by an Enhanced Neural Network. *J. Mol. Biol*. **214**:171–182.

Levin, J.M., Robson, B., and Garnier, J. (1986) An algorithm for secondary structure determination in proteins based on sequence similarity. *Febs Lett*., **205**(2):303–308.

McClelland, J.L., and Rumelhart, D.E. (1988) *Explorations in Parallel Distributed Processing* vol 3. pp 318–362. MIT Press, Cambridge MA.

Rost, B. (1996). PHD: predicting one-dimensional protein structure by profile based neural networks. *Meth. in Enzym*., **266**:525–539.

Stultz, C.M., White, J.V., and Smith, T.F. (1993) Structural Analysis Based on State-space Modeling. *Prot. Science*, **2**:305–314.

White, J.V., Stultz, C.M., and Smith, T.F. (1994) Protein Classification by Stochastic Modeling and Optimal Filtering of Amino-Acid Sequences. *Math. Biosc.*, **119**:35–75.

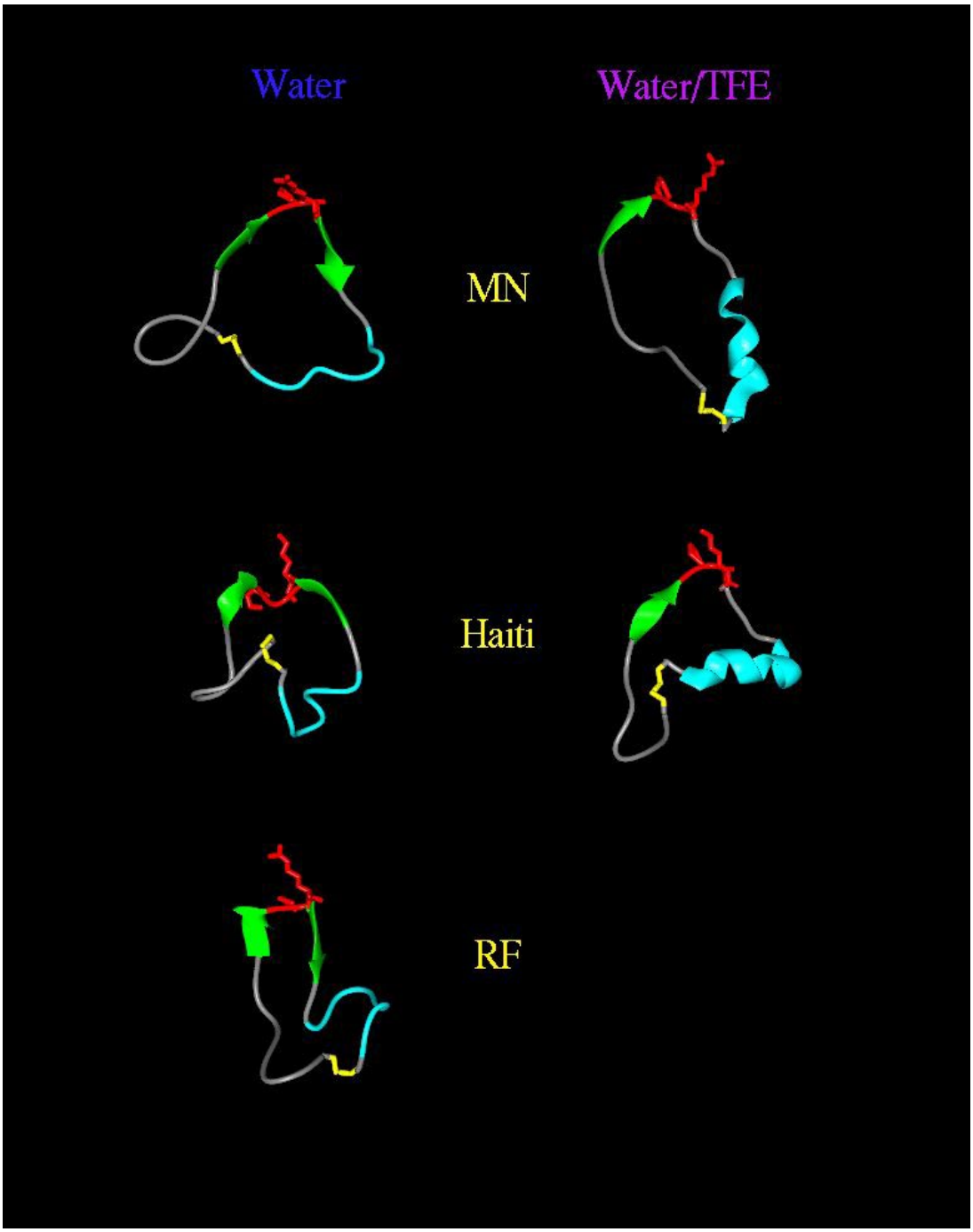

Figure 1. Ribbon diagram showing the average folding patterns of the structures of the MN, Haiti and RF V3 loops in water and in a mixture of 70%/30% water/TFE. In each case the average is done over 70 sampled low energy structures. Note that, in each case, the neutralizing epitope containing the central GPG(R/K) sequence forms a protruding loop even though the local structure and presentation of the loop in the different cases are noticeably different. Structures that satisfy the NMR constraints of the V3 loops in water show a higher degree of flexibility than those in agreement with the NMR data in the mixed water/TFE solvent. This is due to the formation of the alpha helix in the mixed solvent. Color code is as follow: GPG(R/K) crest is red, extended regions are green, disulfide bridges are yellow and the alpha helical region is cyan.

```
                  1       10        20        30
                  |        |         |         |
V3MN              CTRPNYNKRKRIHIGPGRAFYTTKNIIGTIRQAHC
NMR (water)       .EECTTTTCCEEEETTTTEEEECCCCTTTTTTCC.
NMR (TFE)         .EECTTTTEEEECCTTTTCCCCCHHHHHHHHHHC.
PSA method        CCCCCCCCCCEEEECCTCCCCCCCHHHHHHHHHHH
Gibrat method     ECCCCCCCEEEEEEECCCCEEEEEEEHEEEEECCC
Levin method      .TSSSSCSSSCEEECTTCCEEEEECCCCHHHHSSC
DPM method        CCCTTCTCHHHECCTCCCHHCEECCEEEEEHHHCC
SOPMA predict     ECCTTHCTEEEEEEECCCCCCTTTCCCCTTCCCCE
DSC method        CCCCCCCCCEEEEECCCCCEEECCCCCCCCCCCCC
PHD method        CCCCCCCCCEEEE.CCC.EE..CC........CCC
nnpredict method  CCCCCCCCCCCCCCCCCEEEECCCCEEECHHCCCC
                  1       10        20        30
                  |        |         |         |
V3Haiti           CTRPNDNTRKSIPMGPGKAFYATGDIIGNIRQAHC
NMR (water)       .CECTTTTEEECCCTTTTEECCCCCCTTTTTTTC.
NMR (TFE)         .EECTTTTCCEEEETTTTCCTTTTCHHHHHHHHH.
PSA method        CCCCCCCCCCCCCCCCCCCCCCCCCHHCCHHHHHC
Gibrat method     ECCCCCCCCCCCCCCCCCCCEEEEEHHEEEEEECCE
Levin method      TCCCCCCSSCCCCCCSTCCECCTSCCSCSCCHCCC
DPM method        CCCCTCTCCCCCCCTCCCHHCHECCEECCEHHHCC
SOPMA predict     ECCCCTTCCCCETTTCHHHHHETHHHECCCTTCCC
DSC method        CCCCCCCCCCCCCCCCCCCEEECCCCCCCCCCCCCC
PHD method        CCCCCCCC.....CCCC.EE.CCC..........C
nnpredict method  CCCCCCCCCCCCCCCCCCCCEECCCCCEECHHCCCC
                  1       10        20        30
                  |        |         |         |
V3RF              CTRPNNNTRKSITKGPGRVIYATGQIIGDIRKAHC
NMR (water)       .CTTTTEECEECCCTTTTEECCCTTTTCCTTTTC.
PSA method        CCCCCCCCCCCCCCCCCCCEEEECCCHHHHHHHHHH
Gibrat method     ECCCCCCCCEEEEECCCCEEEEEECEECEEEEEHCC
Levin method      CCCCCCCSHS..CCCTTCEEEECCCCCCSCCCCC.
DPM method        CCTTTCTCCCCECCTCCCEEEEECCEECCEHHHCC
SOPMA predict     ECCCCCCCEEEEECCCCCCCEEETTCEEEHHHHHHH
DSC method        CCCCCCCCCCEEECCCCCEEEEECCCCCCCCCCCCC
PHD method        CCCCCCCC..E..CCCC.EEE.CC.EE.......C
nnpredict method  CCCCCCCCCCCCCCCCCCEEEECCCCEEECHHCCCC
```

Figure 2. Comparison of secondary structure assignments of the NMR determined structures and secondary structure prediction for the three V3 loops, MN, Haiti and RF. The different prediction algorithms are indicated on the left. Some of these methods are discussed by Myers and Farmer in Part III of this compendium.

Meaning of Symbols

H	alpha helix	T	turn
C	random coil/loop	S	bend
E	strand	.	unassigned

Key to Prediction Algorithms

PSA	Stultz *et al.*	Gibrat	Gibrat *et al.*
Levin	Levin *et al.*	DPM	Deleage *et al.*
SOPM	Geourjon *et al.*	DSC	King *et al.*
PHD	Rost *et al.*	nnpredict	McClelland *et al.*